

## RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical  
Information Center (STIC) no errors detected.

Application Serial Number: 10/807,228A  
Source: 1FW16  
Date Processed by STIC: 11/22/04

# ***ENTERED***



IFW16

**RAW SEQUENCE LISTING**

DATE: 11/22/2004

PATENT APPLICATION: US/10/807,228A

TIME: 14:32:22

Input Set : N:\Cr3\RULE60\10807228A.RAW

Output Set: N:\CRF4\11222004\J807228A.raw

**SEQUENCE LISTING**

```

3 (1) GENERAL INFORMATION:
5     (i) APPLICANT: Sogabe, Atsushi
6           Hattori, Takashi
7           Nishiya, Yoshiaki
8           Kawamura, Yoshihisa
10    (ii) TITLE OF INVENTION: NOVEL CREATINE AMIDINOHYDROLASE, PRODUCTION
11           THEREOF AND USE THEREOF
13    (iii) NUMBER OF SEQUENCES: 3
15    (iv) CORRESPONDENCE ADDRESS:
16          (A) ADDRESSEE: LEYDIG, VOIT & MAYER, LTD.
17          (B) STREET: Two Prudential Plaza, Suite 4900
18          (C) CITY: Chicago
19          (D) STATE: Illinois
20          (E) COUNTRY: US
21          (F) ZIP: 60601-6780
23    (v) COMPUTER READABLE FORM:
24          (A) MEDIUM TYPE: Floppy disk
25          (B) COMPUTER: IBM PC compatible
26          (C) OPERATING SYSTEM: PC-DOS/MS-DOS
27          (D) SOFTWARE: PatentIn Release #1.0, Version #1.25
29    (vi) CURRENT APPLICATION DATA:
C--> 30          (A) APPLICATION NUMBER: US/10/807,228A
C--> 31          (B) FILING DATE: 23-Mar-2004
32          (C) CLASSIFICATION:
34    (vii) PRIOR APPLICATION DATA:
W--> 35          (A) APPLICATION NUMBER: US/08/799,897
36          (B) FILING DATE: 13-FEB-1997
W--> 38          (A) APPLICATION NUMBER: JP 25435/1996
39          (B) FILING DATE: 13-FEB-1996
41    (viii) ATTORNEY/AGENT INFORMATION:
42          (A) NAME: Robert F. Green
43          (B) REGISTRATION NUMBER: 27555
44          (C) REFERENCE/DOCKET NUMBER: 78064
46    (ix) TELECOMMUNICATION INFORMATION:
47          (A) TELEPHONE: (312) 616-5600
48          (B) TELEFAX: (312) 616-5700
49          (C) TELEX: 25-3533
51 (2) INFORMATION FOR SEQ ID NO: 1:
53     (i) SEQUENCE CHARACTERISTICS:
54          (A) LENGTH: 404 amino acids
55          (B) TYPE: amino acid
56          (D) TOPOLOGY: linear

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/807,228A

DATE: 11/22/2004

TIME: 14:32:22

Input Set : N:\Crif3\RULE60\10807228A.RAW

Output Set: N:\CRF4\11222004\J807228A.raw

W--> 58 (ii) MOLECULE TYPE:

59 (A) DESCRIPTION: protein

61 (vi) ORIGINAL SOURCE:

62 (A) ORGANISM: *Alcaligenes faecalis*

63 (B) STRAIN: TE3581 (FERM P-14237)

65 (ix) FEATURE:

66 (A) NAME/KEY: mat peptide

67 (B) LOCATION: 1 to 404

68 (D) OTHER INFORMATION: protein having creatine amidino-

69 hydrolase activity

71 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

73	Met	Thr	Asp	Asp	Met	Leu	His	Val	Met	Lys	Trp	His	Asn	Gly	Glu	Lys
74	1				5					10					15	
75	Asp	Tyr	Ser	Pro	Phe	Ser	Asp	Ala	Glu	Met	Thr	Arg	Arg	Gln	Asn	Asp
76				20					25					30		
77	Val	Arg	Gly	Trp	Met	Ala	Lys	Asn	Asn	Val	Asp	Ala	Ala	Leu	Phe	Thr
78			35					40					45			
79	Ser	Tyr	His	Cys	Ile	Asn	Tyr	Tyr	Ser	Gly	Trp	Leu	Tyr	Cys	Tyr	Phe
80		50					55				60					
81	Gly	Arg	Lys	Tyr	Gly	Met	Val	Ile	Asp	His	Asn	Asn	Ala	Thr	Thr	Ile
82	65					70				75					80	
83	Ser	Ala	Gly	Ile	Asp	Gly	Gly	Gln	Pro	Trp	Arg	Arg	Ser	Phe	Gly	Asp
84				85						90				95		
85	Asn	Ile	Thr	Tyr	Thr	Asp	Trp	Arg	Arg	Asp	Asn	Phe	Tyr	Arg	Ala	Val
86			100					105					110			
87	Arg	Gln	Leu	Thr	Thr	Gly	Ala	Lys	Arg	Ile	Gly	Ile	Glu	Phe	Asp	His
88		115					120						125			
89	Val	Asn	Leu	Asp	Phe	Arg	Arg	Gln	Leu	Glu	Glu	Ala	Leu	Pro	Gly	Val
90		130					135					140				
91	Glu	Phe	Val	Asp	Ile	Ser	Gln	Pro	Ser	Met	Trp	Met	Arg	Thr	Ile	Lys
92	145				150					155					160	
93	Ser	Leu	Glu	Glu	Gln	Lys	Leu	Ile	Arg	Glu	Gly	Ala	Arg	Val	Cys	Asp
94				165						170					175	
95	Val	Gly	Gly	Ala	Ala	Cys	Ala	Ala	Ala	Ile	Lys	Ala	Gly	Val	Pro	Glu
96			180					185					190			
97	His	Glu	Val	Ala	Ile	Ala	Thr	Thr	Asn	Ala	Met	Ile	Arg	Glu	Ile	Ala
98		195					200						205			
99	Lys	Ser	Phe	Pro	Phe	Val	Glu	Leu	Met	Asp	Thr	Trp	Thr	Trp	Phe	Gln
100		210					215						220			
101	Ser	Gly	Ile	Asn	Thr	Asp	Gly	Ala	His	Asn	Pro	Val	Thr	Asn	Arg	Ile
102	225				230					235					240	
103	Val	Gln	Ser	Gly	Asp	Ile	Leu	Ser	Leu	Asn	Thr	Phe	Pro	Met	Ile	Phe
104				245						250					255	
105	Gly	Tyr	Tyr	Thr	Ala	Leu	Glu	Arg	Thr	Leu	Phe	Cys	Asp	His	Val	Asp
106			260					265					270			
107	Asp	Ala	Ser	Leu	Asp	Ile	Trp	Glu	Lys	Asn	Val	Ala	Val	His	Arg	Arg
108		275					280						285			
109	Gly	Leu	Glu	Leu	Ile	Lys	Pro	Gly	Ala	Arg	Cys	Lys	Asp	Ile	Ala	Ile
110		290					295						300			

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/807,228A

DATE: 11/22/2004

TIME: 14:32:22

Input Set : N:\CrF3\RULE60\10807228A.RAW

Output Set: N:\CRF4\11222004\J807228A.raw

```

111 Glu Leu Asn Glu Met Tyr Arg Glu Trp Asp Leu Leu Lys Tyr Arg Ser
112 305                      310                      315                      320
113 Phe Gly Tyr Gly His Ser Phe Gly Val Leu Cys His Tyr Tyr Gly Arg
114                      325                      330                      335
115 Glu Ala Gly Val Glu Leu Arg Glu Asp Ile Asp Thr Glu Leu Lys Pro
116                      340                      345                      350
117 Gly Met Val Val Ser Met Glu Pro Met Val Met Leu Pro Glu Gly Met
118                      355                      360                      365
119 Pro Gly Ala Gly Gly Tyr Arg Glu His Asp Ile Leu Ile Val Gly Glu
120                      370                      375                      380
121 Asp Gly Ala Glu Asn Ile Thr Gly Phe Pro Phe Gly Pro Glu His Asn
122 385                      390                      395                      400
123 Ile Ile Arg Asn
124                      404
126 (2) INFORMATION FOR SEQ ID NO: 2:
128     (i) SEQUENCE CHARACTERISTICS:
129         (A) LENGTH: 1212 base pairs
130         (B) TYPE: nucleic acid
131         (C) STRANDEDNESS: double
132         (D) TOPOLOGY: linear
134     (ii) MOLECULE TYPE: genomic DNA
136     (vi) ORIGINAL SOURCE:
137         (A) ORGANISM: Alcaligenes faecalis
138         (B) STRAIN: TE3581 (FERM P-14237)
140     (ix) FEATURE:
141         (A) NAME/KEY: CDS
142         (B) LOCATION: 1 to 1212
144     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
146 ATG ACT GAC GAC ATG TTG CAC GTG ATG AAA TGG CAC AAC GGC GAG AAA 48
147 Met Thr Asp Asp Met Leu His Val Met Lys Trp His Asn Gly Glu Lys
148 1 5 10 15
149 GAT TAT TCG CCG TTT TCG GAT GCC GAG ATG ACC CGC CGC CAA AAC GAC 96
150 Asp Tyr Ser Pro Phe Ser Asp Ala Glu Met Thr Arg Arg Gln Asn Asp
151 20 25 30
152 GTT CGC GGC TGG ATG GCC AAG AAC AAT GTC GAT GCG GCG CTG TTC ACC 144
153 Val Arg Gly Trp Met Ala Lys Asn Asn Val Asp Ala Ala Leu Phe Thr
154 35 40 45
155 TCT TAT CAC TGC ATC AAC TAC TAT TCC GGC TGG CTG TAC TGC TAT TTC 192
156 Ser Tyr His Cys Ile Asn Tyr Tyr Ser Gly Trp Leu Tyr Cys Tyr Phe
157 50 55 60
158 GGA CGC AAG TAC GGC ATG GTC ATC GAC CAC AAC AAC GCC ACG ACG ATT 240
159 Gly Arg Lys Tyr Gly Met Val Ile Asp His Asn Asn Ala Thr Thr Ile
160 65 70 75 80
161 TCG GCC GGC ATC GAC GGC GGC CAG CCC TGG CGC CGC AGC TTC GGC GAC 288
162 Ser Ala Gly Ile Asp Gly Gly Gln Pro Trp Arg Arg Ser Phe Gly Asp
163 85 90 95
164 AAC ATC ACC TAC ACC GAC TGG CGC CGC GAC AAT TTC TAT CGC GCC GTG 336
165 Asn Ile Thr Tyr Thr Asp Trp Arg Arg Asp Asn Phe Tyr Arg Ala Val
166 100 105 110

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/807,228A

DATE: 11/22/2004

TIME: 14:32:22

Input Set : N:\Crif3\RULE60\10807228A.RAW

Output Set: N:\CRF4\11222004\J807228A.raw

```

167 CGC CAG CTG ACC ACG GGC GCC AAG CGC ATC GGC ATC GAG TTC GAC CAC 384
168 Arg Gln Leu Thr Thr Gly Ala Lys Arg Ile Gly Ile Glu Phe Asp His
169      115      120      125
170 GTC AAT CTC GAC TTC CGC CGC CAG CTC GAG GAA GCC CTA CCG GGC GTC 432
171 Val Asn Leu Asp Phe Arg Arg Gln Leu Glu Glu Ala Leu Pro Gly Val
172      130      135      140
173 GAC TTC GTC GAC ATC AGC CAG CCC TCG ATG TGG ATG CGC ACC ATC AAG 480
174 Glu Phe Val Asp Ile Ser Gln Pro Ser Met Trp Met Arg Thr Ile Lys
175 145      150      155      160
176 TCG CTC GAA GAG CAG AAG CTG ATC CGC GAA GGC GCC CGC GTG TGT GAC 528
177 Ser Leu Glu Glu Gln Lys Leu Ile Arg Glu Gly Ala Arg Val Cys Asp
178      165      170      175
179 GTC GGC GGC GCG GCC TGC GCG GCT GCC ATC AAG GCC GGC GTG CCC GAG 576
180 Val Gly Gly Ala Ala Cys Ala Ala Ala Ile Lys Ala Gly Val Pro Glu
181      180      185      190
182 CAT GAA GTG GCG ATC GCC ACC ACC AAT GCG ATG ATC CGC GAG ATC GCC 624
183 His Glu Val Ala Ile Ala Thr Thr Asn Ala Met Ile Arg Glu Ile Ala
184      195      200      205
185 AAA TCG TTC CCC TTC GTG GAG CTG ATG GAC ACC TGG ACC TGG TTC CAG 672
186 Lys Ser Phe Pro Phe Val Glu Leu Met Asp Thr Trp Thr Trp Phe Gln
187      210      215      220
188 TCG GGC ATC AAC ACC GAC GGC GCG CAC AAT CCG GTC ACC AAC CGC ATC 720
189 Ser Gly Ile Asn Thr Asp Gly Ala His Asn Pro Val Thr Asn Arg Ile
190 225      230      235      240
191 GTG CAA TCC GGC GAC ATC CTT TCG CTC AAC ACC TTC CCG ATG ATC TTC 768
192 Val Gln Ser Gly Asp Ile Leu Ser Leu Asn Thr Phe Pro Met Ile Phe
193      245      250      255
194 GGC TAC TAC ACC GCG CTG GAG CGC ACG CTG TTC TGC GAC CAT GTC GAT 816
195 Gly Tyr Tyr Thr Ala Leu Glu Arg Thr Leu Phe Cys Asp His Val Asp
196      260      265      270
197 GAC GCC AGC CTC GAC ATC TGG GAG AAG AAC GTG GCC GTG CAT CGC CGC 864
198 Asp Ala Ser Leu Asp Ile Trp Glu Lys Asn Val Ala Val His Arg Arg
199      275      280      285
200 GGG CTC GAG CTG ATC AAG CCG GGC GCG CGC TGC AAG GAC ATC GCC ATC 912
201 Gly Leu Glu Leu Ile Lys Pro Gly Ala Arg Cys Lys Asp Ile Ala Ile
202      290      295      300
203 GAG CTC AAC GAG ATG TAC CGC GAG TGG GAC CTG CTG AAG TAC CGC TCC 960
204 Glu Leu Asn Glu Met Tyr Arg Glu Trp Asp Leu Leu Lys Tyr Arg Ser
205 305      310      315      320
206 TTC GGC TAT GGC CAC TCC TTC GGC GTG CTG TGC CAC TAC TAC GGT CGC 1008
207 Phe Gly Tyr Gly His Ser Phe Gly Val Leu Cys His Tyr Tyr Gly Arg
208      325      330      335
209 GAG GCC GGC GTG GAG CTG CGC GAG GAC ATC GAC ACC GAG CTG AAG CCC 1056
210 Glu Ala Gly Val Glu Leu Arg Glu Asp Ile Asp Thr Glu Leu Lys Pro
211      340      345      350
212 GGC ATG GTG GTC TCC ATG GAG CCG ATG GTG ATG CTG CCG GAG GGC ATG 1104
213 Gly Met Val Val Ser Met Glu Pro Met Val Met Leu Pro Glu Gly Met
214      355      360      365
215 CCC GGT GCC GGC GGC TAT CGC GAG CAC GAC ATC CTG ATC GTC GGG GAG 1152

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/807,228A

DATE: 11/22/2004

TIME: 14:32:22

Input Set : N:\Crf3\RULE60\10807228A.RAW

Output Set: N:\CRF4\11222004\J807228A.raw

```

216 Pro Gly Ala Gly Gly Tyr Arg Glu His Asp Ile Leu Ile Val Gly Glu
217      370                      375                      380
218 GAC GGT GCC GAG AAC ATC ACC GGC TTC CCG TTC GGT CCG GAA CAC AAC 1200
219 Asp Gly Ala Glu Asn Ile Thr Gly Phe Pro Phe Gly Pro Glu His Asn
220 385                      390                      395                      400
221 ATC ATC CGC AAC
222 Ile Ile Arg Asn
223      404
226 (2) INFORMATION FOR SEQ ID NO: 3:
228     (i) SEQUENCE CHARACTERISTICS:
229         (A) LENGTH: 39 base pairs
230         (B) TYPE: nucleic acid
231         (C) STRANDEDNESS: single
232         (D) TOPOLOGY: linear
234     (ii) MOLECULE TYPE: other nucleic acid (synthetic DNA)
236     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
238 CAACATGTCG TCAGTCATAT GTGTTTCCTG TGTGAAATT

```

39

## VERIFICATION SUMMARY

PATENT APPLICATION: US/10/807,228A

DATE: 11/22/2004

TIME: 14:32:23

Input Set : N:\Crf3\RULE60\10807228A.RAW

Output Set: N:\CRF4\11222004\J807228A.raw

L:30 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]  
L:31 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]  
L:38 M:238 W: Alpha Fields not Ordered, Reordered [(A) APPLICATION NUMBER:] of (1)(vii)  
L:58 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=1